BULLETIN of the
SCHOOL OF EDUCATION
INDIANA UNIVERSITY

Schoolroom Motivation: I. Two Studies of Quantity and Pattern of Verbal Reinforcement as Related to Performance on a Routine Task

by
ERVIN KAPOS
EDMUND V. MECH
and
WILLIAM H. FOX

Vol. 33, No. 1 January, 1957

OFFICERS OF THE UNIVERSITY

HERMAN B WELLS, A.M., LL.D., President

HERMAN THOMPSON BRISCOE, Ph.D., Vice-President, and Dean of the Faculties

JOSEPH AMOS FRANKLIN, B.S., Vice-President, and Treasurer

JOHN WILLIAM ASHTON, Ph.D., LL.D., Vice-President, and Dean of Student and Educational Services

WENDELL WILLIAM WRIGHT, Ph.D., Vice-President, Dean of the School of Education, and Director of Administrative Studies and Institutional Relations

CHARLES EDWIN HARRELL, A.B., LL.B., Registrar, and Director of the Office of Records and Admissions

ADMINISTRATIVE STAFF OF THE SCHOOL OF EDUCATION

WENDELL WILLIAM WRIGHT, Ph.D., Dean

HOWARD T. BATCHELDER, Ph.D., Associate Dean

PHILIP PEAK, Ph.D., Assistant Dean

(MRS.) DOROTHY J. MCCAW, Assistant to the Dean

OTTO HUGHES, Ed.D., Principal of the University School

KATHLEEN DUGDALE, A.M., Editor of Publications in Education

ADMINISTRATIVE STAFF OF THE DIVISION OF RESEARCH AND FIELD SERVICES

HAROLD H. CHURCH, Ph.D., Director

WILLIAM H. FOX, Ph.D., Assistant Director

RUTH FAY ARGANBRIGHT, M.S. in Ed., Assistant Director

W. MONFORT BARR, Ed.D., Assistant Director of School Surveys

PAUL W. SEAGERS, Ed.D., School Building Consultant

AUGUST W. EBERLE, Ph.D., School Community Consultant

BULLETIN OF THE SCHOOL OF EDUCATION
INDIANA UNIVERSITY

SCHOOLROOM MOTIVATION: I. TWO STUDIES OF QUANTITY AND PATTERN OF VERBAL REINFORCEMENT AS RELATED TO PERFORMANCE ON A ROUTINE TASK

by

ERVIN KAPOS
EDMUND V. MECH
and
WILLIAM H. FOX

Published by

DIVISION OF RESEARCH AND FIELD SERVICES

Indiana University

Price, $1.00, postpaid

For sale by the Indiana University Bookstore

Bloomington, Indiana

Entered as second-class mail matter September 30, 1924, at the post office at Bloomington, Indiana, under the Act of August 24, 1912. Published in January, March, May, July, September, and November of each year, from the University Office, Bloomington, Indiana.

TABLE OF CONTENTS

Chapter		Page
I.	ORIENTATION	1
	The Experiment in Education	1
	The Problem	3
II.	METHODOLOGY	5
	The Variables	5
	The Subjects	6
	The Task	7
	The Procedure	7
	The Data	8
	The Analysis	9
III.	EXPERIMENT ONE: MASSED PRACTICE	11
	Verbal Reinforcement and Intelligence	11
	The Pattern of Verbal Reinforcement	16
	The Quantity of Verbal Reinforcement	20
	Summary	23
IV.	EXPERIMENT TWO: SPACED PRACTICE	25
	Verbal Reinforcement and Intelligence	25
	The Pattern of Verbal Reinforcement	29
	The Quantity of Verbal Reinforcement	32
	Summary	37
V.	FURTHER CONSIDERATIONS	38
	Conclusions	38
	Two Other Important Variables	38
	An Improved Experimental Design	39
	Verbal Reinforcement and Motivation	40
	REFERENCES	42

LIST OF TABLES

Table		Page
1.	Schedules and Grade Levels of the Experimental Groups	6
2.	Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 100% Massed Reinforcement Group	11
3.	Pretest, Experimental, and Adjusted Means for the Three Subgroups of Varying Intelligence Levels Within the 100% Massed Reinforcement Group	12
4.	Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 75% Massed Reinforcement Group	12
5.	Pretest, Experimental, and Adjusted Means for the Three Subgroups of Varying Intelligence Levels Within the 75% Massed Reinforcement Group	13
6.	Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 50% Massed Reinforcement Group	13
7.	Pretest, Experimental, and Adjusted Means for the Three Subgroups of Varying Intelligence Levels Within the 50% Massed Reinforcement Group	14
8.	Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 25% Massed Reinforcement Group	14
9.	Pretest, Experimental, and Adjusted Means for the Three Subgroups of Varying Intelligence Levels Within the 25% Massed Reinforcement Group	15
10.	Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 0% Massed Reinforcement Group	15
11.	Pretest, Experimental, and Adjusted Means for the Three Subgroups of Varying Intelligence Levels Within the 0% Massed Reinforcement Group	16

Table Page

12. Analysis of Covariance for Differences in Performance Between the Groups Receiving 75% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively 17

13. Pretest, Experimental, and Adjusted Means for the Groups Receiving 75% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively . 17

14. Analysis of Covariance for Differences in Performance Between the Groups Receiving 50% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively 18

15. Pretest, Experimental, and Adjusted Means for the Groups Receiving 50% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively . 18

16. Analysis of Covariance for Differences in Performance Between the Groups Receiving 25% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively 19

17. Pretest, Experimental, and Adjusted Means for the Groups Receiving 25% Massed Reinforcement on a Periodic and on a Random Schedule, Respectively . 19

18. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Massed Verbal Reinforcement (Periodic and Random Schedules Pooled) 20

19. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Massed Verbal Reinforcement (Periodic and Random Reinforcement Groups Pooled) 21

20. The t-Tests for Differences Between the Adjusted Means for the Groups Receiving 100%, 75%, 50%, 25%, and 0% Massed Verbal Reinforcement 21

21. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Massed Reinforcement on Periodic Schedules . . 22

22. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Massed Reinforcement on Periodic Schedules 22

23. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Massed Reinforcement on Random Schedules . . . 23

Table | Page

24. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Massed Reinforcement on Random Schedules 23

25. Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 100% Spaced Reinforcement Group 25

26. Pretest, Experimental, and Adjusted Means for the Subgroups of Varying Intelligence Levels Within the 100% Spaced Reinforcement Group 26

27. Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 75% Spaced Reinforcement Group 26

28. Pretest, Experimental, and Adjusted Means for the Subgroups of Varying Intelligence Levels Within the 75% Spaced Reinforcement Group 26

29. Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 50% Spaced Reinforcement Group 27

30. Pretest, Experimental, and Adjusted Means for the Subgroups of Varying Intelligence Levels Within the 50% Spaced Reinforcement Group 27

31. Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 25% Spaced Reinforcement Group 28

32. Pretest, Experimental, and Adjusted Means for the Subgroups of Varying Intelligence Levels Within the 25% Spaced Reinforcement Group 28

33. Analysis of Covariance for Differences in Performance Between Subgroups of Varying Intelligence Levels Within the 0% Spaced Reinforcement Group 29

34. Pretest, Experimental, and Adjusted Means for the Subgroups of Varying Intelligence Levels Within the 0% Spaced Reinforcement Group 29

35. Analysis of Covariance for Differences in Performance Between the Groups Receiving 75% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively 30

Table | Page

36. Pretest, Experimental, and Adjusted Means for the Groups Receiving 75% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively . 30

37. Analysis of Covariance for Differences in Performance Between the Groups Receiving 50% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively 31

38. Pretest, Experimental, and Adjusted Means for the Groups Receiving 50% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively . 31

39. Analysis of Covariance for Differences in Performance Between the Groups Receiving 25% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively 32

40. Pretest, Experimental, and Adjusted Means for the Groups Receiving 25% Spaced Reinforcement on a Periodic and on a Random Schedule, Respectively . 32

41. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Spaced Verbal Reinforcement (Periodic and Random Schedules Pooled) 33

42. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Spaced Verbal Reinforcement (Periodic and Random Reinforcement Groups Pooled) 33

43. The t-Tests for Differences Between the Adjusted Means for Groups Receiving 100%, 75%, 50%, 25%, and 0% Spaced Verbal Reinforcement 34

44. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Spaced Reinforcement on Periodic Schedules . . 34

45. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Spaced Reinforcement on Periodic Schedules 35

46. The t-Tests for Differences Between the Adjusted Means for Groups Receiving 100%, 75%, 50%, 25%, and 0% Spaced Verbal Reinforcement on Periodic Schedules 35

47. Analysis of Covariance for Differences in Performance Between Groups Receiving Varying Quantities of Spaced Reinforcement on Random Schedules . . . 36

Table Page

48. Pretest, Experimental, and Adjusted Means for the Groups Receiving Varying Quantities of Spaced Reinforcement on Random Schedules 36

49. The t-Tests for Differences Between the Adjusted Means for Groups Receiving 100%, 75%, 50%, 25%, and 0% Spaced Reinforcement on Random Schedules . 37

ACKNOWLEDGMENT

The authors are deeply indebted to Wendell W. Wright, Vice-President of Indiana University and Dean of the School of Education, for his continuing support of basic educational experimentation.

The authors also wish to acknowledge the whole-hearted assistance, in making all the necessary initial arrangements for the experiments, of Homer H. Terrell, Principal of the Central School in Martinsville, Indiana; and of Luther A. Lockwood, Superintendent of Schools, and Miss L. Frances Smith, Director of Elementary Education in Columbus, Indiana. Thanks are also due the principals of the participating schools, including, in addition to Terrell, the following: Mrs. Frieda Crider (North School, Martinsville), Miss Beulah G. Hastings (South School, Martinsville), Lloyd Reece (McKinley School, Columbus), Herbert Reese Twenty-Fifth Street School, Columbus), Robert L. Snyder (Farm Shop School, Martinsville), and Ben Walker (State Street School, Columbus).

The success of educational experimentation in classroom situations is clearly dependent to a large extent upon the cooperation of the classroom teacher. The authors wish to express their gratitude to the following teachers for the courtesy and highly competent cooperation they received: Mrs. Dorothy Asher, Mrs. Doris Cowgil, Mrs. Barbara Cummins, Mrs. Louise Hartley, Mrs. Naomi Kanter, Mrs. Donna Lozar, Mrs. Eunice O'Brien, and Mrs. Ruth Rusie, all of the Central School, Martinsville; Mrs. Gladys Buckles, of the North School, Martinsville; Mrs. Dea Curtis and Mrs. Louise Davee of the South School, Martinsville; David Rowland and Mrs. Vera Roderick of the Farm Shop School, Martinsville; Miss Marilyn Romine and Miss Martha Ruddick of the McKinley School, Columbus; Mrs. Doris Fulp and Mrs. Ruth Neal of the State Street School, Columbus; and Miss Thelma Davis of the Twenty-Fifth Street School, Columbus.

E. K.
E. V. M.
W. H. F.

CHAPTER I

ORIENTATION

The following is a report of two experiments, carried out in 1954 and 1955, to discover how elementary school pupils react to positive verbal stimuli from their regular classroom teachers. The two experiments were identical in design, except that the first dealt with the effects of positive verbal stimuli (which will be referred to as verbal reinforcement in what follows) under massed practice, while the second dealt with the effects of the same stimuli under spaced practice. The intertrial interval under massed practice was essentially zero, whereas there was an interval of some 24 hours between trials under spaced practice.

The Experiment in Education

The authors of the present report have felt it desirable to present what might be described as a "justification" of the experimental approach, as typified by the studies reported here, to problems in the field of education. In their view, this justification must rest essentially on the interdependence of educational practice and educational research (specifically, the experiment) in formulating a rigorous science of education.

Most educators are well aware of the fact that the problems confronting education have not fallen easily and rapidly in the face of the "scientific" movement in education. Yet, frustrating as this lack of tangible progress has been, there has been continuous progress in formulating new and more comprehensive theories and in accumulating basic information. The question now confronting the educator and researcher therefore appears to be the following: Can this progress be utilized in tackling the problems of education, or must such research and theory construction be condemned as "academic"?

The success and meaningful utilization of basic educational research and theory construction depends critically upon the active participation of the only possible link between theory and research and actual classroom practice: the classroom teacher and the school administrator.

Theory construction in education, while arising primarily out of what is generally called basic research, is essentially a two-way process. First, if basic educational research and theory construction are to have any reason for existing, the inquiries and methods of the researcher must be motivated by the problems of the classroom. This, in turn, implies that the educators in direct day-to-day contact with these problems must apprise the researcher not only of the problems, but also of the shortcomings of existing tools and methods in tackling

them. This they can hardly be expected to do unless they can be convinced that the final solution of such problems must be achieved primarily by basic research.

Second, the new results of basic research must be submitted by the researchers to the acid test of the actual classroom situation. Clearly, this again requires the cooperation of the classroom teacher and the school administrator, and their conviction of the fundamental "worthwhileness" of the cooperative effort involved.

Once the theory has been critically examined in the light of reality, and possibly even "tried out," it may not prove adequate to explain or predict fully the observed phenomena. Even if it does, it must be revised continuously in the light of fresh facts relevant to it. How critically important, therefore, is the open-minded collaboration of the classroom teacher and the educational researcher and theorist!

A brief examination of the major problem underlying the present studies will give some indication of the importance of the experimental method in educational research. It is desired to discover the effect of varying quantities of praise upon school-age children in a general situation. Once this problem has been thoroughly investigated it will presumably be possible to determine, for any given situation, what amount of praise will lead to optimum learning, or performance.

There are at least four different ways in which this problem can be attacked. First, an attempt may be made to provide an <u>a priori</u> answer from purely theoretical grounds. Yet, with so many imponderables involved in the behavior of the human organism, no such answer is likely to cover the situation sufficiently completely, and the process of adjusting such an answer to fit the facts may prove only a little less tiresome than an ordinary "trial and error" procedure.

Second, the problem may be approached by querying a number of classroom teachers who may be expected to be acutely aware of the problem and of its practical context. But their replies may be expected to be quite imprecise, and often couched in terms of special instances. There is no way of determining to what extent the teacher's response is based on behavioral observation and to what extent it is influenced by existing theory. Above all, the replies will be to a large extent determined by the individual characteristics of the teachers, and will not reflect systematic observation under carefully controlled and standardized conditions.

Third, the researcher may attack the problem through personal observation. This might remove some of the biases that his data would have if collected from a large number of classroom teachers, but it will introduce the researcher's personal bias. The researcher must still evaluate to what extent each phrase of the teacher constitutes praise in the classroom situation, and he must still attempt to gauge learning or performance from observation. Clearly, the observational errors in such a procedure would cast doubt over any results obtained.

It would thus appear that, if it is to provide results of reasonably assured validity, any procedure employed in approaching the problem must involve the administration of objectively quantifiable reinforcement under controlled and standardized conditions, and must make provision for the objective, quantitative measurement of resulting response changes. These specifications describe exactly the fourth possible method of approach: the experimental method.

In his recent evaluation of the relative productivity of various research techniques, Cornell (8:33) states: "Despite objections to it, the experiment will undoubtedly continue as one of the most important techniques in educational research." Elsewhere in the same report he says: "All types [of research] covered in this review are essential and important in educational research. In the judgment of the writer, however, greater contributions to building a science of education may come from the rigorous methods of experimentation."

It therefore appears that, if we are to discover what causes changes in pupil behavior and if we are to relate these discoveries to the operations a teacher performs in the classroom, we have no recourse but to carry out rigorous experimental research in the classroom. Though the classroom situation can never be as completely controlled as a laboratory situation, such experimentation need not be the less rigorous for that.

If basic, theoretical experimentation is to be valuable, it must be conducted under the most critical standards of scientific rigor. By the same token, it must be communicated to school personnel at all levels to permit, and even invite, appraisal in the classroom situation. The present authors believe that these two requirements are not incompatible.

The Problem

The question of reinforcement and motivation is quite properly considered a facet of the theory of learning, and the present experiments may therefore properly be referred to as "learning" studies. The techniques relative to the routine task and the verbal reinforcement were similar to those described by Hurlock (16) in an earlier study, and to those employed by Auble, Mech, and others (1, 2, 3, 4, 5, 17, 18) in a series of exploratory studies preceding the present experiments. These techniques are more fully described in succeeding sections.

Of the numerous questions still open in this generally neglected field, the present experiments were designed to cast some light on the following:

1. Do varying quantities of verbal reinforcement, as defined for these studies, exert a differential effect upon pupil behavior, as measured by performance on a routine task?

2. Does the presence or absence of a regular, cyclic pattern of reinforcement exert a differential effect upon the effectiveness of a fixed quantity of verbal reinforcement?

3. Does a fixed quantity of verbal reinforcement exert a differential effect upon pupils of sharply varying intelligence levels?

4. To what extent do results obtained under massed practice correspond to those obtained under spaced practice?

The above questions have been stated in practical terminology for convenience, but the sophisticated reader can easily restate the first three in the form of null hypotheses.

How important are these questions in terms of classroom practice, especially in view of the restricted definitions for "verbal reinforcement" and "pupil behavior" adopted in the interest of scientific rigor?

The authors believe that, if the questions can be answered under the restrictive definitions adopted, it will be relatively easy, if tedious, to extend the results to the limits of their generalizability. It is difficult to visualize any more profitable approach to the demands of rigorous theory construction, in spite of the tedium involved in part of the work.

The authors do not subscribe to the notion that the motivation of pupils can be explained solely in terms of the verbal stimuli emitted by the teacher. But these overt "artificial" incentives do constitute one manipulable aspect of motivation, and therefore provide a particularly suitable area for basic experimental research, the results of which it is within the classroom teacher's power to apply directly.

For a cogent appraisal of the merits of various other incentives, the reader is referred to Stephens (21). Studies in this area have produced few clear-cut results.

CHAPTER II

METHODOLOGY

The Variables

The principal independent variable in the present experiments is readily identified as the quantity of verbal reinforcement administered to the pupils in a classroom. This variable is a discrete one, taking on only five distinct values: 100%, 75%, 50%, 25%, and 0%. Evidently the "100% reinforcement" group was the group receiving verbal reinforcement after every trial, while the "0% reinforcement" group served as a control group. In the case of the "partial reinforcement" groups (75%, 50%, and 25%) the total group was reinforced after the indicated proportion of the trials.

The other independent variable is the pattern of verbal reinforcement. This is also a discrete variable, taking on only two distinct values: periodic and random. Clearly this distinction applies only to the groups receiving partial reinforcement, as the other groups are reinforced after all and none of the trials, respectively. Under periodic partial reinforcement the indicated proportion of reinforcement is delivered according to a fixed cyclic pattern throughout the experiment. In the case of the corresponding random partial reinforcement, verbal reinforcement is administered after the same number of trials, but these trials are randomly selected.

The variable dependent upon the foregoing independent variables is pupil performance on a routine task. Both the task and the measure of performance are described in succeeding sections.

A further variable, which can be called neither dependent nor independent, but should rather be identified as a criterion of classification, was the intelligence level of the subjects. The measures of intelligence employed were the Otis Intelligence Test and the California Test of Mental Maturity, but all subjects under any given experimental condition had the same measure. The intelligence quotients for the subjects were drawn from the cumulative school records. Subjects were then divided into "high," "medium," and "low" intelligence groups on the basis of these intelligence quotients as follows: high, 110 or more; medium, 90 to 109; low, 89 or less. One reason for this choice of boundaries for the intelligence groups lies in certain considerations of comparability between the two different measures of intelligence that had to be used.

The assignment to the various experimental conditions of the 10 classrooms used in each experiment is summarized in Table 1. The two experiments were precisely identical with respect to these aspects of the design.

TABLE 1. SCHEDULES AND GRADE LEVELS OF THE EXPERIMENTAL GROUPS

Group	Reinforcement schedule		Grade-level
	Quantity	Pattern	
A_1	100%	Periodic-random	4
A_2	100%	Periodic-random	3
B_1	75%	Periodic	4
B_2	75%	Random	4
C_1	50%	Periodic	4
C_2	50%	Random	3
D_1	25%	Periodic	3
D_2	25%	Random	3
E_1	0%	Periodic-random	4
E_2	0%	Periodic-random	3

The Subjects

The subjects used in the two experiments described here were third grade and fourth grade pupils* drawn from the elementary school systems of Martinsville and Columbus, Indiana.

Each of the experiments involved five third grade classrooms and five fourth grade classrooms, giving a total of 20 classrooms used in the two studies. The assignment of these classrooms to the various experimental groups is indicated in Table 1.

A subject was included in the final tabulation and analysis of the data only if "complete" data were obtained for the subject. This was defined to include (a) a measure of I.Q., (b) at least two of the three pretest trials, and (c) at least 22 of the 28 experimental trials, with no four consecutive trials missing.

This criterion yielded 325 complete sets of data for the experiment involving massed practice, and 284 sets for the experiment involving spaced practice.

*Concerning the choice of subjects at these grade levels, it may be noted that Gesell and Ilg (13) describe the child at 8 years as one who "loves praise and to be reminded of his improvement." (p. 414)

The Task

The routine arithmetical task confronting the subjects on each trial consisted of a set of additions and subtractions which were prepared as follows.

Each set consisted of two pages of problems; each page had nine rows and each row had eight problems. There were thus 144 problems in each set, and it would clearly have been impossible for any third or fourth grade pupil to complete the problems within the four-minute time limit established for each trial. The maximum score was, in fact, under 100.

All problems involved only the addition or subtraction of pairs of numbers of, at most, two digits. All numbers were selected by drawing successive pairs of digits from a table of random integers (11). Clearly, if the first digit of a pair was zero, the resulting number was less than 10.

The choice of operation for each pair of numbers was made as follows: First, all number pairs where the second number was larger than the first were designated as addition problems. If there were more than 36 such number pairs, any over 36 were replaced by randomly selected number pairs where the second number was less than the first. If there were less than 36 such number pairs, additional ones were designated as addition problems randomly, until there were 36 addition problems. The remaining number pairs were designated as subtraction problems. Thus the only non-random limitation imposed was that each page of 72 problems should finally contain 36 addition problems and 36 subtraction problems.

Five forms of such sets of problems were administered to the subjects in cyclic order. Thus, over the three pretest trials and 28 experimental trials, four of the five forms appeared six times and one appeared seven times.

The Procedure

Each of the two experiments consisted of three pretest trials (which provided a measure of initial performance) and of 28 experimental trials. Each trial was of precisely four minutes' duration.

In the case of the experiment dealing with massed practice, the pretest trials were administered on the afternoon of the first day of the experiment, and the experimental trials were administered on the morning of the second day. Each subject was handed a set of 28 experimental forms, and was required to keep the form for any given trial covered until the trial began. The interval between trials sufficed merely to administer the verbal reinforcement, if any, and to prepare for the next trial. There was a rest interval of approximately 10 minutes after the fourteenth experimental trial.

The experiment dealing with spaced practice extended over a period of eight weeks. The pretest trials were administered during the first three days of the first week, and were followed by an interval of one and one-half weeks. The experimental trials, which were administered at the rate of one per day, about mid-morning, from Monday to Friday, began on Monday during the third week. No trial was administered on Friday of the fifth week (after the fourteenth experimental trial) to provide a longer interval corresponding to the rest interval allowed at that stage during the experiment dealing with massed practice.

Prior to the beginning of the experiments the teachers were presented with complete sets of instructions for carrying out the study (including a schedule of reinforcements), and their questions were carefully answered by one of the experimenters. In this way the entire conduct of the experiment in each classroom was in the hands of the regular teacher, and the experimenters did not at any time enter any of the classrooms.

The teachers were requested to avoid delivering reinforcement in any form during the pretest trials, and to remain stationary at the front of the room. During the experimental trials, however, all of the teachers were required to move about the room constantly, pretending to note how each pupil was performing.*

If verbal reinforcement was scheduled after a given trial, the teacher was requested to convey the following four points in her own words, as was the case with all the experimental instructions: (a) "You all did very well this time." (b) "I (the teacher) am very pleased." (c) "Now, try to do even better on the next set of problems." (d) "Work as fast as you can, but try not to make any mistakes."

The Data

As indicated, the data concerning the measures of intelligence administered to all subjects were drawn from the cumulative school records. Raw scores were not recorded for the purposes of the present experiments; only intelligence quotients were used.

The sets of problems administered in the course of the pretest and experimental trials were scored by hand. There were 10,075 such sets for the first experiment (three pretest and 28 experimental trials for 325 subjects), and 8,804 sets (284 subjects) for the second experiment. All raw data were

---

*According to the observations of one of the experimenters, this procedure was not adequately followed in only one of the 20 classrooms--the class receiving 50% periodic reinforcement under massed practice.

recorded on summary sheets. Subjects under each experimental condition were immediately separated into "high," "medium," and "low" intelligence groups. Less than one per cent of the raw data were missing due to pupil absence, etc.

The measure of initial performance recorded for each subject in both experiments was the mean number of problems solved correctly during the three pretest trials. The corresponding experimental measure was the number of problems solved correctly during the 28 experimental trials.

The Analysis

The statistical technique employed to analyze the data resulting from these experiments was the analysis of covariance with a single criterion of classification--the experimental condition. The pretest and experimental measures have already been described in the section on the data.

The analysis of covariance was originally defined under the assumption of equal or proportional group frequencies, and the analysis in the present case was somewhat complicated by the necessarily unequal number of subjects in the various groups. A general solution of the analysis of covariance for unequal groups has been derived by Tsao (23), but Edwards and Parkin (9) report that the usual method of the analysis of covariance is not greatly affected by a moderate degree of disproportionality in the number of cases per group. Gourlay (14), moreover, points out that any bias due to Snedecor's (20) method is only increased by the application of Tsao's modification. It was therefore decided to employ the usual technique of the analysis of covariance, as described by Snedecor.

The analyses of the two sets of experimental data proceed exactly parallel, through the following three steps:

1. An analysis for the differential effects of a fixed quantity of verbal reinforcement upon subjects of varying intelligence levels.

2. An analysis for the differential effects of a fixed quantity of verbal reinforcement administered according to varying patterns.

3. An analysis for the differential effects of varying quantities of verbal reinforcement.

It should be pointed out that a comparison could not be carried out between the data obtained under massed practice and the data obtained under spaced practice, because the subjects who participated in the experiment utilizing spaced practice were learning arithmetic continuously during the eight weeks spanned by the study. Thus, at least part of any gain made by them, as compared with that made by the subjects who participated in the study utilizing massed practice, would

be due not to the differential effect of comparable quantities of verbal reinforcement administered under massed and spaced practice but simply to school learning experiences.

CHAPTER III

EXPERIMENT ONE: MASSED PRACTICE

Verbal Reinforcement and Intelligence

As described earlier, the subjects in every reinforcement group were divided into "high," "medium," and "low" intelligence subgroups according to whether their I.Q.'s (as measured by the Otis Intelligence Test or the California Test of Mental Maturity) were 110 or more, 90 to 109, or 89 or less, respectively. Though the choice of these boundaries for the groups was dictated by certain considerations of comparability between the two tests, the boundaries had the further advantage that 110 and 90 were approximately the upper and lower quartiles for most of the experimental groups.

The hypothesis tested for any given reinforcement group is essentially that a fixed amount of reinforcement exerts no differential effects on subjects of varying intelligence levels. No cross-comparisons between reinforcement groups are involved here.

Table 2 summarizes the analysis of covariance carried out to test for differences in performance between the subgroups of high, medium, and low intelligence levels within the 100% reinforcement group. The obtained F-value of 0.29, for two degrees of freedom between groups and 61 degrees of freedom within groups, indicates quite clearly that the null hypothesis must be retained. It must therefore be concluded that full reinforcement exerts essentially the same effect upon subjects of varying intelligence levels when the level of initial performance is taken into account.

TABLE 2. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 100% MASSED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	63	1,693.91		
Within groups	61	1,677.74	27.50	
Between adjusted means	2	16.17	8.09	0.29

Table 3 presents the pretest, experimental, and adjusted means for the groups compared in the analysis of covariance summarized in Table 2. The last of these means is obtained from the experimental mean by an adjustment involving the pretest mean and the regression of the experimental measure on the pretest measure. The process is clearly described by Garrett (12). Inspection of Table 3 indicates that the small obtained value of F was not due to any exceptional variation within groups, but rather to the fact that the adjusted means were quite close together.

TABLE 3. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE THREE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 100% MASSED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	16	17.96	20.92	19.96
Medium	33	18.64	22.05	20.26
Low	16	13.35	16.63	21.29

Table 4 summarizes a similar analysis of covariance carried out for groups of varying intelligence levels within the 75% reinforcement group. As in the case of all other analyses in this section, the groups receiving 75% reinforcement on periodic and random schedules were pooled. The obtained value of F was 0.06, which fell far short of statistical significance.

TABLE 4. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 75% MASSED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	72	2,398.44		
Within groups	70	2,394.35	34.21	
Between adjusted means	2	4.09	2.05	0.06

Table 5 presents the pretest, experimental, and adjusted means associated with the foregoing analysis. Inspection of

the table again indicates that the differences between the adjusted means were negligible.

TABLE 5. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE THREE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 75% MASSED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	11	17.55	21.55	25.81
Medium	48	24.17	29.63	26.41
Low	15	14.98	18.85	26.01

A similar analysis of covariance for differences between subgroups of varying intelligence levels was carried out for the groups receiving 50% reinforcement, and this analysis is summarized in Table 6. The obtained F-value of 0.04 was clearly not significant.

TABLE 6. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 50% MASSED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	61	1,345.54		
Within groups	59	1,343.56	22.77	
Between adjusted means	2	1.98	0.99	0.04

Inspection of the pretest, experimental, and adjusted means associated with this analysis, which are presented in Table 7, again shows only negligible differences between the adjusted means. It may be noted that there was a fairly sharp difference between the pretest and experimental means for the high and medium intelligence subgroups on the one hand, and those for the low intelligence group on the other. Therefore the retention of the null hypothesis implies not that no differences in performance existed, but, rather, that the performances of the three subgroups relative to initial performance were virtually identical.

TABLE 7. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE THREE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 50% MASSED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	19	18.26	19.56	19.23
Medium	34	18.65	19.84	19.10
Low	10	14.90	16.46	19.62

Table 8 summarizes the analysis of covariance for differences in performance between the subgroups of differing intelligence levels within the 25% reinforcement group. Though the obtained F-value of 2.45 was considerably higher than the F-values obtained in the preceding analyses, it falls well short of significance at the 0.05 level of confidence, with two degrees of freedom between groups and 64 degrees of freedom within groups.

TABLE 8. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 25% MASSED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	66	577.37		
Within groups	64	536.31	8.38	
Between adjusted means	2	41.06	20.53	2.45

Inspection of Table 9, which presents the pretest, experimental, and adjusted means associated with the foregoing analysis, shows a difference of some two points between the adjusted mean for the high intelligence group and those for the medium and low intelligence groups. A correspondingly high difference may be observed in the pretest and experimental means. The 25% reinforcement group was the only one exhibiting this peculiarity. (Table 5 shows that the pretest and experimental means for the medium intelligence subgroup within the 75% reinforcement group were considerably higher than the corresponding means for the subgroups of high and low intelligence levels.) Inspection of the data fails to

suggest that the distribution of I.Q.'s in the medium intelligence subgroup of this group differed in any essential way from the corresponding distribution for other groups.

TABLE 9. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE THREE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 25% MASSED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	12	14.44	18.67	14.42
Medium	36	10.56	12.15	12.32
Low	20	8.75	10.54	12.77

Though there can, of course, be no question concerning the differential effects of reinforcement upon the subgroups of varying intelligence levels within the 0% group, a similar analysis had to be computed for this group. The analysis of covariance is summarized in Table 10, and the associated pretest, experimental, and adjusted means are presented in Table 11.

TABLE 10. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 0% MASSED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	53	1,502.45		
Within groups	51	1,498.02	29.37	
Between adjusted means	2	4.43	2.22	0.08

The obtained value of F was 0.08, and all differences between adjusted means were negligible, even though there were considerable differences between the pretest and experimental means.

TABLE 11. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE THREE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 0% MASSED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	18	25.87	29.10	24.07
Medium	25	21.57	23.97	23.46
Low	12	12.86	15.18	23.81

It must therefore be concluded that, though pupils of varying levels of intelligence make very largely uniform gains (when initial performance is considered) simply as a result of practice, when no reinforcement is administered, they also make uniform relative gains under partial and total reinforcement. No relationship therefore appears to exist between the effect of a given quantity of reinforcement upon a pupil and the level of intelligence of the pupil.*

The Pattern of Verbal Reinforcement

The pattern of a fixed quantity of partial verbal reinforcement has been defined as the sequence of reinforced and unreinforced trials. A reinforcement pattern may belong to one of two major classes: systematic or random. Of the large number of possible systematic patterns for partial reinforcement, only cyclic patterns of period 4 have been chosen for use in the present experiment. Thus, since the quantities of partial reinforcement are expressed in terms of fourths (75%, 50%, 25%), a random sequence of four reinforced and unreinforced trials was determined for each partial reinforcement group, and this cycle of four was repeated seven times over the 28 experimental trials.

Analyses of covariance were computed for each of the partial reinforcement groups between subgroups receiving periodic and random reinforcement, respectively. The analysis for the 75% reinforcement group is summarized in Table 12. The F-value obtained was 15.51, which is significant considerably

*It must be pointed out that this conclusion could not be drawn without the evidence furnished by Tables 10 and 11. It would be possible, for instance, for the adjusted performance of one subgroup under 0% reinforcement to lag significantly behind the other two. If, then, the adjusted performances under partial and total reinforcement were uniform, the above conclusion could not be drawn.

beyond the 0.01 level of confidence for one degree of freedom between groups and 71 degrees of freedom within groups.

TABLE 12. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 75% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	72	2,398.44		
Within groups	71	1,967.30	27.71	
Between adjusted means	1	431.14	431.14	15.51*

*Significant at the 0.01 level of confidence.

Inspection of Table 13, which presents the pretest, experimental, and adjusted means for the groups compared in the foregoing analysis, indicates that the adjusted mean for the group receiving 75% reinforcement on a periodic schedule was markedly higher than that for the group receiving the same quantity of reinforcement on a random schedule. Since the variable under consideration was actually experimentally controlled and manipulated in this instance, it is possible to consider this an indication of a possible causal relationship between the variation in pattern of reinforcement and the variation in performance.

TABLE 13. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 75% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group*	n	Pretest mean	Experimental mean	Adjusted mean
75% P	34	21.59	29.17	28.86
75% R	40	21.09	23.76	24.02

*75% P denotes the group receiving 75% reinforcement on a periodic schedule, while 75% R denotes the group receiving the same amount of reinforcement on a random schedule. This notation is to be similarly interpreted in all succeeding tables.

A similar analysis of covariance was computed for differences in performance between the subgroups of the 50% reinforcement group receiving their reinforcement on a periodic and on a random schedule, respectively. The results of this analysis are summarized in Table 14. It is readily seen that the obtained value of F (3.41) is no longer statistically significant at the 0.05 level of confidence. However, application of the method described by Burke (6) indicates that this F-value is significant at the 0.08 level of confidence for one degree of freedom between groups and 60 degrees of freedom within groups. (Cochran and Cox (7) present cogent arguments against rigidly ignoring the results of an analysis of covariance unless significant at the 0.05 level.)

TABLE 14. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 50% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	61	1,345.54		
Within groups	60	1,273.22	21.22	
Between adjusted means	1	72.32	72.32	3.41*

*Significant at the 0.08 level of confidence.

Inspection of Table 15, which presents the pretest, experimental, and adjusted means for the groups being compared above, indicates that the group receiving 50% reinforcement on a random schedule was superior to that receiving the same amount of reinforcement on a periodic schedule. In this connection, the reader is referred to the note on page 8 of this report.

TABLE 15. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 50% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group	n	Pretest mean	Experimental mean	Adjusted mean
50% P	35	13.57	13.84	18.37
50% R	28	23.39	25.95	20.27

A similar analysis of covariance for differences in performance between the subgroups receiving 25% reinforcement on a periodic and those receiving it on a random schedule, respectively, is summarized in Table 16. The obtained value of F was 2.46, and this is found to be significant only at the 0.12 level of confidence for one degree of freedom between groups and 65 degrees of freedom within groups. This means that the probability of an error of the first kind would be 0.12 if the null hypothesis were rejected. It is therefore preferable not to do so, but merely to consider this result, in the context of the two preceding ones, as indicative of a possible difference in adjusted performance due to varying patterns of reinforcement.

TABLE 16. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 25% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	66	577.37		
Within groups	65	556.30	8.56	
Between adjusted means	1	21.07	21.07	2.46*

*Significant at the 0.12 level of confidence.

Inspection of the table of pretest, experimental, and adjusted means associated with the foregoing analysis (Table 17) indicates that the adjusted performance of the subgroup receiving periodic reinforcement is superior to that of the subgroup receiving random reinforcement.

TABLE 17. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 25% MASSED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group	n	Pretest mean	Experimental mean	Adjusted mean
25% P	36	10.79	13.43	13.35
25% R	32	10.64	12.15	12.24

It has thus been demonstrated that certain differences exist between groups receiving partial reinforcement on periodic and random schedules. These differences were significant at the 0.01 level of confidence for the 75% reinforcement group, at the 0.08 level for the 50% group, and at the 0.12 level for the 25% group. Thus the probability of an error of the first kind in rejecting the null hypothesis increases as the quantity of reinforcement decreases. Though the evidence is not clear-cut, it appears that the differences that exist are in favor of the groups receiving reinforcement on a periodic schedule.

The Quantity of Verbal Reinforcement

The principal independent variable in the present experiment was the quantity of reinforcement administered in the course of the series of 28 trials. Two classrooms were assigned randomly to each of the 100%, 75%, 50%, 25%, and 0% reinforcement groups. (In the case of the partial reinforcement groups, the total group was reinforced after the indicated proportion of trials.)

An analysis of covariance was computed for differences in performance between the five reinforcement groups. Inspection of Table 18 indicates that the obtained value of F was 4.10, which is significant beyond the 0.01 level of confidence. It may be concluded, therefore, that significant differences in performance on a routine task do exist between groups receiving varying quantities of reinforcement.

TABLE 18. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF MASSED VERBAL REINFORCEMENT (PERIODIC AND RANDOM SCHEDULES POOLED)

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	323	8,068.74		
Within groups	319	7,674.06	24.06	
Between adjusted means	4	394.68	98.67	4.10*

*Significant at the 0.01 level of confidence.

Inspection of Table 19, which presents the pretest, experimental, and adjusted means for the five reinforcement groups, indicates that the highest adjusted mean is that for

the 75% reinforcement group. This is followed by the 100%, 25%, 0%, and 50% groups, in that order.

TABLE 19. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF MASSED VERBAL REINFORCEMENT (PERIODIC AND RANDOM REINFORCEMENT GROUPS POOLED)

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	65	17.17	20.44	20.90
75%	74	21.32	26.25	22.06
50%	63	17.94	19.22	18.82
25%	68	10.72	12.83	20.51
0%	55	21.08	23.73	19.81

To determine precisely which of the experimental conditions exerted differential effects upon adjusted performance, the experimental groups were compared pairwise by means of the t-test for the analysis of covariance described by Cochran and Cox (7). The approximation for the error of the difference proposed by Finney (10) was employed, since the number of degrees of freedom within groups was sufficiently large.

Table 20 presents the results of these t-tests. In evaluating differences between the 50% reinforcement group and other groups, the reader is again referred to the note on page 8 of this report.

TABLE 20. THE t-TESTS FOR DIFFERENCES BETWEEN THE ADJUSTED MEANS FOR THE GROUPS RECEIVING 100%, 75%, 50%, 25%, and 0% MASSED VERBAL REINFORCEMENT

Group	75%	50%	25%	0%
100%	1.14	2.24*	0.45	1.19
75%		3.78**	1.85***	2.53*
50%			1.93***	1.07
25%				0.77

*Significant at the 0.05 level of confidence.

**Significant at the 0.01 level of confidence.

***Significant at the 0.10 level of confidence.

Since the preceding portions of the analysis suggested that differing patterns of reinforcement might be accountable for variations in adjusted performance, separate analyses of covariance were computed for differences in adjusted performance between the groups receiving reinforcement on periodic schedules and for those between groups receiving reinforcement on random schedules. Clearly, the 100% and 0% reinforcement groups appeared unchanged in both cases.

The analysis of covariance computed for groups receiving periodic reinforcement is summarized in Table 21. The obtained value of F was 7.83, which is significant considerably beyond the 0.01 level of confidence for four degrees of freedom between groups and 219 degrees within groups.

TABLE 21. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF MASSED REINFORCEMENT ON PERIODIC SCHEDULES

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	223	5,971.47		
Within groups	219	5,224.14	23.85	
Between adjusted means	4	747.33	186.83	7.83*

*Significant at the 0.01 level of confidence.

Inspection of the adjusted means, presented in Table 22, indicates essentially the same differences as those observed in the analysis for pooled groups.

TABLE 22. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF MASSED REINFORCEMENT ON PERIODIC SCHEDULES

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	65	17.17	20.44	20.49
75% P	34	21.59	29.17	24.09
50% P	35	13.57	13.84	18.06
25% P	36	10.79	13.43	20.98
0%	55	21.08	23.73	19.24

The analysis of covariance for differences in performance between groups receiving varying quantities of reinforcement on random schedules is summarized in Table 23. The obtained value of F was only 0.57, indicating that the null hypothesis must be retained in this instance.

TABLE 23. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF MASSED REINFORCEMENT ON RANDOM SCHEDULES

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	218	5,316.51		
Within groups	214	5,260.89	24.58	
Between adjusted means	4	55.62	13.91	0.57

Table 24 presents the associated pretest, experimental, and ajusted means.

TABLE 24. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF MASSED REINFORCEMENT ON RANDOM SCHEDULES

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	65	17.17	20.44	22.11
75% R	40	21.09	23.76	21.15
50% R	28	23.39	25.95	20.84
25% R	32	10.64	12.15	20.94
0%	55	21.08	23.73	21.14

Summary

Three hundred and twenty-five subjects, drawn from five third grade and five fourth grade classrooms, were tested on a routine arithmetical task for three pretest and 28 massed experimental trials. Two classes received each of the following quantities of reinforcement: 100%, 75%, 50%, 25%, and 0%; of

each pair of classes one was reinforced on a fixed periodic schedule and the other on a predetermined random schedule. Subjects in each group were classified into high, medium, and low intelligence subgroups.

Analyses of covariance between the subgroups of differing intelligence levels within each reinforcement group indicate the absence of any relationship between intelligence and the effect of a fixed quantity of reinforcement upon adjusted performance.

Similar analyses between the subgroups of each of the partial reinforcement groups receiving periodic and random reinforcement revealed a difference significant at the 0.01 level of confidence in the 75% group, a difference significant at the 0.08 level in the 50% group, and a difference significant at the 0.13 level in the 25% group. The analyses in the 75% and 25% groups indicated that the subgroups reinforced on periodic schedules performed better than those reinforced on random schedules.

The analysis of covariance for differences in performance between groups receiving varying quantities of reinforcement resulted in an F-value significant at the 0.01 level of confidence. The highest adjusted mean was that for the 75% reinforcement group, followed, in order, by the 100%, 25%, 0%, and 50% groups.

Separate analyses indicated that significant differences existed between groups receiving periodic reinforcement but not between those receiving random reinforcement.

CHAPTER IV

EXPERIMENT TWO: SPACED PRACTICE

Verbal Reinforcement and Intelligence

As the design for this experiment differed from that for the preceding one only in that the intertrial interval was not less than 24 hours, the analysis to be described is in all respects parallel to the analysis described in Chapter III.

The analysis of covariance for differences in performance between the subgroups of high, medium, and low intelligence levels within the 100% reinforcement group is summarized in Table 25. It is clear from the obtained F-value of 0.85 that no significant differences existed.

TABLE 25. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 100% SPACED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	48	1,086.89		
Within groups	46	1,048.35	22.79	
Between adjusted means	2	38.54	19.27	0.85

Table 26 presents the pretest, experimental, and adjusted means associated with this analysis. Once again, in spite of the sharp divergences between the pretest means, the uniformity of the adjusted means is evident.

Table 27 summarizes the analysis of covariance for differences between the high, medium, and low intelligence subgroups within the 75% reinforcement group. The obtained value of F was 1.74, which falls well short of statistical significance.

The associated pretest, experimental, and adjusted means are shown in Table 28. It should be pointed out here that the number of subjects in the high intelligence group was disproportionately low.

TABLE 26. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 100% SPACED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	13	23.85	39.63	26.99
Medium	26	14.83	25.38	25.37
Low	11	4.12	10.04	25.02

TABLE 27. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLEGENCE LEVELS WITHIN THE 75% SPACED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	59	2,448.61		
Within groups	57	2,308.06	40.49	
Between adjusted means	2	140.55	70.28	1.74

TABLE 28. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 75% SPACED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	9	26.11	36.35	30.58
Medium	29	23.14	35.00	33.06
Low	23	17.99	25.25	29.96

A similar analysis of covariance, carried out for differences in performance between the high, medium, and low intelligence subgroups of the 50% reinforcement group, is summarized in Table 29. The obtained value of F was 0.37, which is certainly far short of statistical significance. The pre-

test, experimental, and adjusted means associated with this analysis are presented in Table 30.

TABLE 29. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 50% SPACED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	50	2,433.87		
Within groups	48	2,396.88	49.93	
Between adjusted means	2	36.99	18.50	0.37

TABLE 30. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 50% SPACED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	15	25.58	42.85	29.07
Medium	28	13.65	26.03	28.35
Low	9	3.70	11.11	26.86

An analysis of covariance was also carried out for differences in performance between the high, medium, and low intelligence subgroups within the 25% reinforcement group. Inspection of Table 31, which summarizes the results of this analysis, shows that the value obtained for F was 0.31. This clearly implies that the null hypothesis of no differences between the subgroups of varying intelligence levels must be retained in the case of the 25% reinforcement group.

The pretest, experimental, and adjusted means for the subgroups of the 25% reinforcement group which were compared in the foregoing analysis are presented in Table 32. The number of subjects in the high intelligence subgroup was again quite low.

As pointed out in the discussion of the analogous results for the study involving massed practice, the foregoing results do not indicate definitely that the effects of a fixed

quantity of verbal reinforcement exerts no differential effect upon the performance of subjects of varying levels of intelligence. If, however, the same lack of significant differences between subgroups of varying intelligence levels were found in the case of the 0% reinforcement group, then such a conclusion could well be drawn.

TABLE 31. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 25% SPACED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	53	3,380.05		
Within groups	51	3,338.99	65.47	
Between adjusted means	2	41.06	20.53	0.31

TABLE 32. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 25% SPACED REINFORCEMENT GROUP

Intelligence level	n	Pretest mean	Experimental mean	Adjusted mean
High	7	12.57	25.63	19.72
Medium	25	10.12	20.71	18.28
low	23	5.28	12.86	17.30

An analysis of covariance for differences in performance was therefore computed for the high, medium, and low intelligence subgroups of the 0% reinforcement (or control) group. The results of this analysis are summarized in Table 33, and the obtained value of F is seen to be 0.37, which falls short of statistical significance.

The pretest, experimental, and adjusted means for the subgroups of the 0% reinforcement group involved in the foregoing covariance analysis are presented in Table 34. In this case, the number of subjects in the low intelligence group was quite small.

TABLE 33. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 0% SPACED REINFORCEMENT GROUP

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	64	2,785.97		
Within groups	62	2,753.31	44.41	
Between adjusted means	2	32.66	16.33	0.37

TABLE 34. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE SUBGROUPS OF VARYING INTELLIGENCE LEVELS WITHIN THE 0% SPACED REINFORCEMENT GROUP

Group	n	Pretest mean	Experimental mean	Adjusted mean
High	24	18.25	30.12	27.21
Medium	37	14.86	26.54	27.90
Low	5	12.87	21.41	25.28

It may therefore be concluded that there exists no statistically significant relationship (nothing can be said about any causal relationship here) between the intelligence level of subjects and their performance on a routine arithmetical task under fixed quantities of verbal reinforcement.

The Pattern of Verbal Reinforcement

As in the case of the experiment involving the use of massed practice, it was thought desirable to determine whether varying patterns of partial reinforcement (where the quantity of reinforcement was fixed) produced differential effects upon adjusted performance on the routine arithmetical task.

Table 35 summarizes the analysis of covariance for differences in performance between groups receiving 75% verbal reinforcement on a periodic schedule and on a random schedule, respectively. The obtained value of F was zero, which indicates that the adjusted means for the two groups were identical. This is borne out by the data presented in Table 36.

TABLE 35. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 75% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	59	2,448.61		
Within groups	58	2,448.61	42.22	
Between adjusted means	1	0	0	0

It should be pointed out that this result does not agree with the result of the corresponding analysis carried out for groups receiving massed reinforcement (Tables 12 and 13). A difference significant at the 0.01 level of confidence was found between the two groups receiving 75% reinforcement under massed practice, with the group that received periodic reinforcement exhibiting the significantly higher adjusted mean score.

TABLE 36. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 75% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group	n	Pretest mean	Experimental mean	Adjusted mean
75% P	32	21.09	30.82	31.53
75% R	29	22.23	32.30	31.53

The analysis of covariance for differences in performance carried out between the groups receiving 50% reinforcement on a periodic schedule and on a random schedule, respectively, is summarized in Table 37.

The obtained value of F is seen to be 11.63, which is significant considerably beyond the 0.01 level of confidence. A similar, though somewhat weaker, difference was observed in the experiment involving the use of massed practice. Inspection of the means for the groups compared in the above analysis which are presented in Table 38, reveals that the difference in this case was in favor of the group receiving reinforcement on a periodic schedule. This result is in agreement with the apparent trend in the previous experiment, but not with the

result for the groups receiving 50% reinforcement under massed practice.

TABLE 37. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 50% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	50	2,433.87		
Within groups	49	1,966.87	40.14	
Between adjusted means	1	467.00	467.00	11.63*

*Significant at the 0.01 level of confidence.

TABLE 38. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 50% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group	n	Pretest mean	Experimental mean	Adjusted mean
50% P	32	6.86	18.02	29.51
50% R	20	28.98	44.74	26.37

Table 39 presents the summary of the analysis of covariance between the groups receiving 25% reinforcement on a periodic schedule and on a random schedule, respectively. The obtained value of F is seen to be 10.97, which is also significant considerably beyond the 0.01 level of confidence.

In this instance, however, inspection of the table of pretest, experimental, and adjusted means (Table 40) indicates that the higher adjusted mean score was that of the group reinforced on a random schedule. This is contrary to the corresponding earlier result.

It has thus been shown that, when partial reinforcement is administered after spaced practice, the varying patterns of reinforcement produced differential effects on performance when the quantity of reinforcement was 50% or 25%. In the former case the higher adjusted mean score was that of the

group receiving periodic reinforcement, while in the latter case the higher adjusted mean score was that of the group receiving random reinforcement. In the case of the groups receiving 75% reinforcement, the two adjusted means were identical.

TABLE 39. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN THE GROUPS RECEIVING 25% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	53	3,380.05		
Within groups	52	2,791.15	53.68	
Between adjusted means	1	588.90	588.90	10.97*

*Significant at the 0.01 level of confidence.

TABLE 40. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING 25% SPACED REINFORCEMENT ON A PERIODIC AND ON A RANDOM SCHEDULE, RESPECTIVELY

Group	n	Pretest mean	Experimental mean	Adjusted mean
25% P	24	10.46	17.44	14.53
25% R	31	6.82	18.53	20.79

There has not been any substantial measure of agreement between the foregoing results and the corresponding results for the group receiving massed reinforcement, except that some significant differences were observed in both cases.

The Quantity of Verbal Reinforcement

To evaluate the effect upon performance of the principal independent variable in this study, an analysis of covariance was computed for differences in performance between the five groups receiving varying quantities of reinforcement.

The partial reinforcement groups included both the groups receiving periodic reinforcement and those receiving random reinforcement. The F-value obtained by this analysis, which is summarized in Table 41, was 4.71, which is significant at the 0.01 level of confidence, for four degrees of freedom between groups and 278 within groups.

TABLE 41. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF SPACED VERBAL REINFORCEMENT (PERIODIC AND RANDOM SCHEDULES POOLED)

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	282	13,080.87		
Within groups	278	12,250.11	44.07	
Between adjusted means	4	830.76	207.69	4.71*

*Significant at the 0.01 level of confidence.

The pretest, experimental, and adjusted means for the five reinforcement groups are presented in Table 42. Inspection of this table shows that the lowest adjusted mean was that for the 75% group, followed by the means for the 100%, 0%, 25% and 50% groups, in that order. These results clearly diverge considerably from those obtained under massed practice.

TABLE 42. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF SPACED VERBAL REINFORCEMENT (PERIODIC AND RANDOM REINFORCEMENT GROUPS POOLED)

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	50	14.82	25.71	26.46
75%	61	21.63	31.52	23.48
50%	52	15.37	28.30	28.34
25%	55	8.41	18.05	27.07
0%	66	15.94	27.49	26.79

As in the case of the preceding study, t-tests were computed for all pairs of reinforcement groups. The results of these tests are presented in Table 43, which shows that the 75% reinforcement group differed from all other groups, but that no other groups differed from each other.

TABLE 43. THE t-TESTS FOR DIFFERENCES BETWEEN THE ADJUSTED MEANS FOR GROUPS RECEIVING 100%, 75%, 50%, 25%, AND 0% SPACED VERBAL REINFORCEMENT

Group	75%	50%	25%	0%
100%	2.29*	1.39	0.46	0.26
75%		3.80*	2.85*	2.74*
50%			0.97	0.43
25%				0.23

*Significant at the 0.05 level of confidence.
*Significant at the 0.01 level of confidence.

Since the analysis summarized in Tables 35 to 40 gives some indication that the pattern of verbal reinforcement plays a significant role in determining the performance of subjects on the routine arithmetical task, analyses similar to the foregoing one were computed separately for groups receiving periodic reinforcement and for groups receiving random reinforcement. Clearly, the 100% and 0% reinforcement groups were unchanged in both of these analyses.

Table 44 summarizes the analysis of covariance for differences in performance between the groups receiving varying quantities of reinforcement on periodic schedules. The obtained value of F was 4.29, which is significant at the 0.01 level of confidence, for four degrees of freedom between groups and 198 within groups.

TABLE 44. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF SPACED REINFORCEMENT ON PERIODIC SCHEDULES

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	202	9,540.00		
Within groups	198	8,778.44	44.34	
Between adjusted means	4	761.56	190.39	4.29*

*Significant at the 0.01 level of confidence.

Inspection of the adjusted means for the groups receiving varying quantities of reinforcement on periodic schedules, which are presented in Table 45, indicates the possible presence of significant differences between pairs of groups other than those noted in Table 43 for periodic and random reinforcement groups pooled.

TABLE 45. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF SPACED REINFORCEMENT ON PERIODIC SCHEDULES

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	50	14.82	25.71	25.16
75% P	32	21.09	30.82	22.06
50% P	32	6.86	18.02	27.90
25% P	24	10.46	17.44	22.60
0%	66	15.94	27.49	29.51

Therefore t-tests were computed for differences between the adjusted means of all pairs of periodically reinforced groups. The results of these t-tests are presented in Table 46.

TABLE 46. THE t-TESTS FOR DIFFERENCES BETWEEN THE ADJUSTED MEANS FOR GROUPS RECEIVING 100%, 75%, 50%, 25%, and 0% SPACED VERBAL REINFORCEMENT ON PERIODIC SCHEDULES

Group	75%	50%	25%	0%
100%	1.99†	1.76‡	1.50	3.37*
75% P		3.40*	0.29	5.03*
50% P			2.85*	1.09
25% P				4.21*

*Significant at the 0.01 level of confidence.
†Significant at the 0.05 level of confidence.
‡Significant at the 0.10 level of confidence.

Table 47 summarizes the analysis of covariance computed for differences in performance between groups receiving varying quantities of verbal reinforcement on predetermined random schedules. The obtained value of F was 5.19, which is significant beyond the 0.01 level of confidence, for four degrees of freedom between groups and 190 within groups.

Inspection of the adjusted means for the randomly reinforced groups presented in Table 48 again suggests that the significant differences between pairs of groups were different from those in the analyses for pooled periodically and randomly reinforced groups and for periodically reinforced groups alone.

TABLE 47. ANALYSIS OF COVARIANCE FOR DIFFERENCES IN PERFORMANCE BETWEEN GROUPS RECEIVING VARYING QUANTITIES OF SPACED REINFORCEMENT ON RANDOM SCHEDULES

Source of variation	d.f.	Sum of squares of adjusted scores	Mean square	F
Total	194	7,482.48		
Within groups	190	6,745.52	35.50	
Between adjusted means	4	736.96	184.24	5.19*

*Significant at the 0.01 level of confidence.

TABLE 48. PRETEST, EXPERIMENTAL, AND ADJUSTED MEANS FOR THE GROUPS RECEIVING VARYING QUANTITIES OF SPACED REINFORCEMENT ON RANDOM SCHEDULES

Group	n	Pretest mean	Experimental mean	Adjusted mean
100%	50	14.82	25.71	27.87
75% R	29	22.23	32.30	24.75
50% R	20	28.98	44.74	28.35
25% R	31	6.82	18.53	31.17
0%	66	15.94	27.49	26.80

As in the preceding analyses, t-tests for differences between the adjusted means of all pairs of randomly reinforced groups were computed. The resulting values of t are presented in Table 49.

From the foregoing analyses it may be concluded that varying quantities of reinforcement do exert differential effects upon performance on a routine arithmetical task under spaced pratice. The precise nature of these differential effects is not comparable to those exerted by the identical quantities of reinforcement under massed reinforcement.

TABLE 49. THE T-TESTS FOR DIFFERENCES BETWEEN THE ADJUSTED MEANS FOR GROUPS RECEIVING 100%, 75%, 50%, 25%, and 0% SPACED REINFORCEMENT ON RANDOM SCHEDULES

Group	75%	50%	25%	0%
100%	2.14*	0.29	2.29*	0.91
75% R		1.98**	3.96*	1.46
50% R			1.57	0.97
25% R				3.19*

*Significant at the 0.01 level of confidence.
**Significant at the 0.05 level of confidence.
***Significant at the 0.10 level of confidence.

Summary

Two hundred and eighty-four subjects, drawn from five third grade and five fourth grade classrooms, were tested on a routine arithmetical task for three pretest and 28 spaced (intertrial interval: 24 hours) experimental trials. Two classes received each of the following quantities of reinforcement: 100%, 75%, 50%, 25%, and 0%. Of each pair of classes, one was reinforced on a fixed periodic schedule and the other on a predetermined random schedule. Subjects in each group were classified into high, medium, and low intelligence subgroups.

Analyses of covariance between the subgroups of differing intelligence levels within each reinforcement group clearly indicate the absence of any relationship between intelligence and the effect of a fixed quantity of reinforcement upon adjusted performance.

Similar analyses between the subgroups receiving periodic and random reinforcement within each of the groups receiving partial reinforcement revealed identical adjusted means for the two subgroups of the 75% reinforcement group, but showed differences significant beyond the 0.01 level of confidence for the 50% and 25% reinforcement groups. In the case of the 50% group, the periodically reinforced subgroup performed better than the randomly reinforced one, but the situation was reversed in the case of the 25% reinforcement group.

The analysis of covariance for differences in performance between groups receiving varying quantities of reinforcement resulted in F-values significant at the 0.01 level of confidence for periodically reinforced groups, randomly reinforced groups, and pooled groups. The pattern of pairwise differences between the groups, however, differed in each of the three analyses.

CHAPTER V

FURTHER CONSIDERATIONS

Conclusions

Four major conclusions may be drawn from the data of the experiments just described.

First, contrary to a notion that has some prevalence in educational circles, the same quantity of verbal reinforcement administered to pupils of sharply varying intelligence levels fails to produce any differential effects related to intelligence upon performance on a routine arithmetical task, provided initial performance is taken into account. It must be stressed that this statement refers only to performance (in terms of proficiency), and should not be lightly extended to say that a fixed quantity of verbal reinforcement produces no differential effects of any sort on children of varying levels of intelligence.

Second, varying patterns of verbal reinforcement do exert differential effects upon performance on a routine arithmetical task. Only two different patterns have been investigated in these studies: a fixed periodic, or cyclic, pattern (of period 4) and a predetermined random pattern. There is no clear indication as to which pattern leads to superior performance.

Third, varying quantities of verbal reinforcement do produce highly significant differential effects upon performance on a routine arithmetical task. The data obtained from these experiments do not indicate an optimum quantity of reinforcement, since the observed differences were not sufficiently systematic.

Fourth, the foregoing major conclusions are equally valid for the experiment using massed practice and for the experiment using spaced practice, though the analyses of the two experiments revealed occasional divergences in specific results.

Two Other Important Variables

In the experiments described here, no account was taken of any teacher-connected variables, such as the personality of the teacher, although it has often been argued that the effectiveness of the incentives administered depends upon such variables.

Though there is unquestionable merit to this argument, it must be noted that one of the primary purposes of the present studies was to investigate the effects of verbal reinforcement administered in a normal classroom situation by the regular classroom teacher. This justifies the random choice of classrooms and teachers.

It must be noted, however, that certain variations in the data of these studies can hardly be explained by any of the experimental procedures. It is entirely possible that some of these variations may be properly ascribed to teacher-connected variables in the classrooms employed in the studies. In view of this, it seems probable that identification of the personality traits which determine the effectiveness of verbal incentives must precede the determination of optimum levels of reinforcement.

A further variable of legitimate experimental interest is delay of verbal reinforcement. In the present experiments all reinforcements were administered immediately upon conclusion of the trials. The question still remains whether varying delays of reinforcement will exert differential effects upon performance on a routine task (where the quantity of reinforcement is held constant).

An Improved Experimental Design

The situation utilized in the experiments described corresponds to the actual schoolroom situation in that the task and the reinforcement were administered by the regular classroom teacher, and that every effort was made to blend the experimental procedure into the everyday routine of the class.

Nevertheless, the situation is artificial in two crucial respects: (a) in using a routine (non-verbal) task for the experimental measure, and (b) in administering the verbal reinforcement to the class as a group, rather than to individuals. When these shortcomings are remedied, the results obtained from such experiments may be expected to have a far greater measure of general validity.

An experimental design utilizing verbal reinforcement administered to individual pupils is presented here, leaving the construction of more complex tasks (both verbal and non-verbal) suitable for experimental use to later research.

The mean number of correct responses for any reinforcement group under the 28 trial design employed in the studies described has been found to be adequately represented by a polynomial curve of the form

$$f(x) = \sum a_i x^i$$

where $i = 1, 2, \ldots, 11$, and the a_i are rational.

It is assumed that the performance curve for any group represents adequately the performance curve for any individual in the group, in the sense that

$$\frac{df_G(x)}{dx} \sim \frac{df_I(x)}{dx}$$

for $1 \leq x \leq 28$, i.e., that the individual curve is at least generally parallel to the group curve at all points, though it may lie above or below it.

A proper reinforcement for a subject is defined as one delivered after a trial when the subject has performed better than during the preceding trial; conversely, an improper reinforcement is one delivered after the subject has done less well than during the preceding trial. A reinforcement after trial x^* is consequently defined to be improper if $\frac{df}{dx}(x^*) < 0$, and proper if $\frac{df}{dx}(x^*) > 0$. (The case when $\frac{df}{dx} = 0$ is easily treated by use of the second derivative.)

It is now possible to select any finite number of combinations in the natural order of, say, 14 (in the case of 50% reinforcement) of the trials 1 through 28, such that the number of improper reinforcements is minimal. Each such combination serves as the reinforcement schedule of one pupil in a class receiving 50% reinforcement. In other respects the experimental procedure remains unchanged.

Essentially, this design permits individual reinforcement to proceed according to a predetermined schedule and thus to be experimentally controlled, with only a minimum possibility that verbal reinforcement will be administered to an individual subject when he has been doing relatively poorly.

The first question to be answered must be whether such reinforcement is significantly more or less effective than group reinforcement. After that come such questions as the studies described here have attempted to answer in the case of group reinforcement.

Verbal Reinforcement and Motivation

In a discussion of Hull's (15) theory of drive, Taylor (22) points out that except in simple experimental arrangements, the performance level of high-drive subjects should be greater than that of low-drive groups, but that higher-drive levels should not always lead to superior performance, in the sense of a greater frequency of occurrence for correct responses.

Recognizing, therefore, that performance, as measured by the number of correct responses, is not necessarily the best measure of the motivation induced by verbal reinforcement, the present authors propose to publish shortly the results of anal-

yses for the two experiments described which will be precisely analogous to the analyses already presented, but which will involve the use of the number of problems attempted in the course of each trial rather than the number correctly solved.

REFERENCES

1. Auble, Donavon, and Mech, Edmund V., "Quantitative Studies of Verbal Reinforcement in Classroom Situations: I. Differential Reinforcement Related to the Frequency of Error and Correct Responses," The Journal of Psychology 35:307-312, April, 1953.

2. Auble, Donavon, and Mech, Edmund V., "Partial Verbal Reinforcement Related to Distributed Practice in a Classroom Situation," The Journal of Psychology 36:165-186, July, 1953.

3. Auble, Donavon, and Mech, Edmund V., "Response Strength in a Classroom Task Related to a 'Forward' Delay in Reinforcement," The Journal of Educational Psychology 45:175-181, March, 1954.

4. Auble, Donavon, and Mech, Edmund V., "Quantitative Studies of Verbal Reinforcement in Classroom Situations: II. Occurrence of Error and Correct Responses Related to a Change in Stimulus Conditions," The Journal of Genetic Psychology 86:207-216, June, 1955.

5. Auble, Donavon; Hurst, Francis M.; and Mech, Edmund V., "Partial Reinforcement Related to Massed and Spaced Training in a Classroom Situation," The Journal of General Psychology (in press).

6. Burke, Cletus J., "Computation of the Level of Significance in the F-test," Psychological Bulletin 48:392-397, September, 1951.

7. Cochran, William G., and Cox, Gertrude M., Experimental Designs, John Wiley, New York, 1950, 454 pp.

8. Cornell, Francis G., "Productive Techniques in Research," Phi Delta Kappan 35:29-34, October, 1953.

9. Edwards, Daisy S., and Parkin, Sidney J., "An Empirical Investigation of the Problem of Disproportionate Frequencies in Analysis of Covariance as Applied to a Methods Experiment," Journal of Experimental Education 22:257-264, March, 1954.

10. Finney, D. J., "Standard Errors of Yields Adjusted for Regression on an Independent Measurement," Biometrics Bulletin 2:53-55, February, 1946.

11. Fisher, Ronald A., and Yates, Frank, Statistical Tables for Biological, Agricultural and Medical Research, Hafner Publishing Company, Inc., New York, 1953, 4th ed., 112 pp.

12. Garrett, Henry E., Statistics in Psychology and Education, Longmans, Green and Co., New York, 1954, 4th ed, 480 pp.

13. Gesell, Arnold L., and Ilg, Frances L., The Child from Five to Ten, Harper and Brothers, New York, 1946, 475 pp.

14. Gourlay, Neil, "Covariance Analysis and Its Applications in Psychological Research," The British Journal of Statistical Psychology 6(pt. 1):25-34, May, 1953.

15. Hull, Clark L., Principles of Behavior, Appleton-Century, New York, 1943, 422 pp.

16. Hurlock, Elizabeth B., "The Evaluation of Certain Incentives Used in School Work," Journal of Educational Psychology 16:145-159, March, 1925.

17. Mech, Edmund V.; Hurst, Francis M.; Auble, Donavon; and Fattu, Nicholas A., An Experimental Analysis of Patterns of Differential Verbal Reinforcement in Classroom Situations, Bulletin of the School of Education, vol. 29, no. 5, Indiana University, Bloomington, September, 1953, 26 pp.

18. Mech, Edmund V.; Kapos, Ervin; Hurst, Francis M.; and Auble, Donavon, "A Mathematical Description of the Relationship Between Frequency of Responding (F) and Intertrial Interval in a Motivational Situation," The Journal of Psychology 37:251-256a, April, 1954.

19. Pearson, Karl, editor, Tables of the Incomplete Beta-function, Biometrika, London, 1948.

20. Snedecor, George W., Statistical Methods, Iowa State College Press, Ames, Iowa, 1946, 4th ed., 485 pp.

21. Stephens, John M., Educational Psychology, Henry Holt, New York, 1951.

22. Taylor, Janet A., "Drive Theory and Manifest Anxiety," Psychological Bulletin 53:303-320, July, 1956

23. Tsao, Fei, "General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," Psychometrika 11:107-128, June, 1946.

BULLETINS IN THE FIELD OF EDUCATION
INDIANA UNIVERSITY

The Bulletin of the School of Education, Indiana University, was first published in 1924 and has appeared regularly since that time. A complete list of the bulletins may be obtained from the School of Education upon request. The studies included in the present and in the two volumes immediately preceding it in date are listed below. Unless otherwise indicated, these may be obtained for $1 each from the Indiana University Bookstore, Bloomington, Indiana.

Volume 31 (1955)

1. Eberle, August William, A Brief History and Analysis of the Operation of the Educational Placement Service at Indiana University, 30 pp.

2. Lynch, William W., Jr., An Approach to the Study of Motivational Problems in Education, 48 pp.

3. Proceedings of the Third Statewide Conference on the Fifth Year in Teacher Education, 69 pp.

4. Watson, Carlos M., and Richey, Robert W., Present Practices and Trends in the Preparation of Elementary School Principals at the Graduate Level, 54 pp.

5. Indiana and Midwest School Building Planning Conference: Proceedings, 149 pp.

6. Crawford, Mary M.; Steinkamp, Stanley; and Hauswald, Edward L., Student Spending at Indiana University, 1951-1952, 82 pp.

Volume 32 (1956)

1. Hoppe, Arthur, Students Help Improve the Curriculum in Indiana, 55 pp.

2. Schmidt, Louis G., Two-Year Evaluation of the Internship in Guidance Program at Indiana University, 67 pp.

3. Strom, Ingrid M., Teaching Load of Teachers of English in Indiana, 63 pp.

4. Hughes, Otto, and Schooler, Virgil E., A Survey of Athletics in the Secondary Schools of the North Central Association, 52 pp.

5. Indiana and Midwest School Building Planning Conference: Proceedings, 58 pp.

6. Barr, W. Monfort; Church, Harold H.; and McGhehey, Marion A., Trends in School District Reorganization in Indiana, 116 pp.

Volume 33 (1957)

1. Kapos, Ervin; Mech, Edmund V.; and Fox, William H., Schoolroom Motivation: I. Two Studies of Quantity and Pattern of Verbal Reinforcement As Related to Performance on a Routine Task, 45 pp.

INDIANA UNIVERSITY
MAIL RECEIVED
UNIVERSITY SCHOOL
BLOOMINGTON, IND.